Promoting reproducibility in addiction research

Marcus R. Munafò [1,2]

1. MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, United Kingdom
2. UK Centre for Tobacco and Alcohol Studies, School of Experimental Psychology, University of Bristol, Bristol, United Kingdom

Corresponding author: Marcus R. Munafò, School of Experimental Psychology, University of Bristol, 12a Priory Road, Bristol BS8 1TU, United Kingdom. E: marcus.munafo@bristol.ac.uk; T: +44.117.9546841; F: +44.117.9288588

Promoting reproducibility in addiction research

In recent years an energetic reproducibility debate has questioned whether published scientific findings are sufficiently robust. The field of addiction science should embrace this debate, and consider measures to improve the quality of the research it conducts. A manifesto has been published that suggests a number of possible measures to substantially reduce the bias towards unreliable positive findings. Addiction*, as a key stakeholder already adopts many of the elements of this manifesto, and will continue to strengthen its policies in the future.*

Is there a reproducibility crisis in biomedical science? In 2005 Ioannidis argued that biases in how science is conducted and reported conspire to reduce the likelihood that published findings are reproducible (1). Since then, growing evidence has emerged that many key findings cannot be replicated. In 2011 scientists from the pharmaceutical company Bayer reported that they were only able to replicate ~25% of results published in the journals they looked at (2). Similar results have been reported by Amgen in 2012 (3) and more recently by the Reproducibility Project: Psychology (4).

A number of factors may serve to undermine the robustness of published findings. Small samples size is one (5, 6), and this appears to hold across the biomedical literature generally (7). What has been termed 'herd behaviour' may increase the risk that scientists will converge on an incorrect answer (8), while cognitive biases, such as confirmation bias and hindsight bias, together with a desire to be recognized, and secure research funding are also potential culprits. Some of these factors may not be directly causal, but instead point to problems with the current incentive structures within which scientists operate. For example, studies conducted in the US, where academic salaries are often not guaranteed if grant income is not generated, tend to over-estimate effects compared to studies

conducted outside the US (9), at least for the "softer" biomedical sciences (10). Studies published in journals with a high Impact Factor seem to be more likely to over-estimate effects (11). Indeed, journal Impact Factor appears to correlate more strongly with the likelihood of retraction than with the number of citations an article receives (12).

INSERT TABLE 1 ABOUT HERE

We recently published a *Manifesto for Reproducible Science* setting out a range of measures to optimize key elements of the scientific process: methods, reporting and dissemination, evaluation and incentives (13). These are shown in Table 1. Many of these measures can be adopted by individual researchers or research groups; others will require the engagement of key stakeholders – funders, journals and institutions. There may also be discipline-specific measures that can be taken – we recently outlined a range of measures intended to improve the reliability of findings generated by functional neuroimaging research (14).

Addiction has introduced a number of measures that align with those described in Table 1. These include encouraging authors to register hypotheses and analysis plans for all types of study (and requiring this for trials), requiring the use of CONSORT and other EQUATOR checklists for trials, encouraging authors to make data sets available, enforcing strict conflict of interest declarations, requiring authors to notify the journal of other articles that have been published or are planned that are based on the same data set as that being reported, and encouraging the use of the Open Science Framework (https://osf.io).

Ultimately, the issue is one of quality control. Science is perhaps like the US automobile industry in the 1970s, where productivity was high but quality control was poor (15). This was the era of the "lemon" – very low quality cars that were essentially built to be fixed later. The Japanese automobile industry took advice from

the US statistician Edwards Deming, and introduced the concept of quality control measures at every stage of the production pipeline. It still has a reputation for reliability today. We need to do something similar in the field of addiction and the journal, Addiction, will be at the forefront of this movement.

References

1. IOANNIDIS, J. P. (2005) Why most published research findings are false, *PLoS Med*, 2, e124.
2. PRINZ, F., SCHLANGE, T. & ASADULLAH, K. (2011) Believe it or not: how much can we rely on published data on potential drug targets?, *Nat Rev Drug Discov*, 10, 712.
3. BEGLEY, C. G. & ELLIS, L. M. (2012) Drug development: Raise standards for preclinical cancer research, *Nature*, 483, 531-3.
4. OPEN-SCIENCE-COLLABORATION (2015) Estimating the reproducibility of psychological science, *Science*, 349, aac4716.
5. BUTTON, K. S., IOANNIDIS, J. P., MOKRYSZ, C. et al. (2013) Power failure: why small sample size undermines the reliability of neuroscience, *Nat Rev Neurosci*, 14, 365-76.
6. HIGGINSON, A. D. & MUNAFO, M. R. (2016) Current Incentives for Scientists Lead to Underpowered Studies with Erroneous Conclusions, *PLoS Biol*, 14, e2000995.
7. DUMAS-MALLET, E., BUTTON, K. S., BORAUD, T., GONON, F. & MUNAFO, M. R. (2017) Low statistical power in biomedical science: a review of three human research domains, *R Soc Open Sci*, 4, 160254.
8. PARK, I. U., PEACEY, M. W. & MUNAFO, M. R. (2014) Modelling the effects of subjective and objective decision making in scientific peer review, *Nature*, 506, 93-6.
9. MUNAFO, M. R., ATTWOOD, A. S. & FLINT, J. (2008) Bias in genetic association studies: effects of research location and resources, *Psychol Med*, 38, 1213-4.
10. FANELLI, D. & IOANNIDIS, J. P. (2013) US studies may overestimate effect sizes in softer research, *Proc Natl Acad Sci U S A*, 110, 15031-6.

11. MUNAFO, M. R., STOTHART, G. & FLINT, J. (2009) Bias in genetic association studies and impact factor, *Mol Psychiatry*, 14, 119-20.

12. BREMBS, B., BUTTON, K. & MUNAFO, M. (2013) Deep impact: unintended consequences of journal rank, *Front Hum Neurosci*, 7, 291.

13. MUNAFO, M. R., NOSEK, B. A., BISHOP, D. V. M. et al. (2017) A manifesto for reproducible science, *Nat Hum Behav*, 1, 0021.

14. POLDRACK, R. A., BAKER, C. I., DURNEZ, J. et al. (2017) Scanning the horizon: towards transparent and reproducible neuroimaging research, *Nat Rev Neurosci*, 18, 115-126.

15. MUNAFO, M., NOBLE, S., BROWNE, W. J. et al. (2014) Scientific rigor and the art of motorcycle maintenance, *Nat Biotechnol*, 32, 871-3.

Table 1. A Manifesto for Reproducible Science.

Theme	Proposal	Examples of initiatives/potential solutions (extent of current adoption)	Stakeholder(s)
Methods	Protecting against cognitive biases	All of the initiatives listed below (* to ****) Blinding (**)	J, F
	Improving methodological training	Rigorous statistical and research methods training of future researchers (*) Rigorous statistical and methods continuing education of researchers (*)	I, F
	Independent methodological support	Involvement of methodologists in research (**) Independent oversight (*)	F
	Collaboration and team science	Multi-site studies / distributed data collection (*) Team science consortia (*)	I, F
Reporting and Dissemination	Promoting study pre-registration	Registered Reports (*) Open Science Framework (*)	J, F
	Improving the quality of reporting	Use of reporting checklists (**) Protocol checklists (*)	J
	Protecting against conflicts of interest	Disclosure of conflicts of interest (***) Exclusion/containment of financial and non-financial conflicts of interest (*)	J
Reproducibility	Transparency and open science	Open data, materials, software, etc. (* to **) Pre-registration (**** for clinical trials, * for other studies)	J, F, R
Evaluation	Diversifying peer review	Preprints (* in biomedical / behavioural sciences, **** in physical sciences) Pre- and post-publication peer-review, e.g. Publons, PubMed Commons (*)	J
Incentives	Reward open, reproducible practices	Badges (*) Registered Reports (*) Transparency and Openness Promotion guidelines (*) Funding replication studies (*) Open science practices in hiring and promotion (*)	J, I, F

Estimated extent of current adoption: * <5%, ** 5-30%, *** 30-60%, **** >60%; Abbreviations for key stakeholders: J: journals / publishers; I: institutions, F: funders, R: regulators. Reproduced with permission from *Nature Human Behaviour* (14).